

BRIDGING THE OUTCOME GAP BETWEEN PROGRAM EFFECTS AND SUBSTANTIVE IMPACTS

**Using the "Likely To" Approach to Extend the Reach of
Evaluation to Substantive and Distant Impacts
American Evaluation Association Conference
November 8th 2003**

**Andy Rowe
GHK International**

OVERVIEW

- ➡ This presentation is based on my reflections on difficult evaluation challenges I have encountered in environment/resource programs and international development
- ➡ I suggest that evaluator statements that desired program effects are *likely to* occur should be accepted as the best we can do, and credible, so long as good enough methods have been employed and only until the programs have a better knowledge base
- ➡ I identify two credible approaches and suggest some initial necessary methodological standards for these

EXAMPLE PROGRAMS FOR THE DISCUSSION

➔ Salmonid Enhancement Program

- Cleaning up fish passages, curtailing excessive harvesting, stocking rivers with hatchery raised young fish

➔ Environmental Conflict Resolution

- Collaborative development of an agreement to resolve an environmental dispute facilitated by a neutral third party

➔ International Peacemaking/keeping and Conflict Resolution

- Multiple forms of interventions (e.g peacekeeping, security sector reform, conflict resolution) to restore necessary conditions for stability and development

➔ Urban Environmental Service Improvement

- Provision of services to address sewage, water supply and quality, roads, pathways and street lighting and similar services in poor communities

➔ Community Development – economic development effects of modest community investments

- Community agencies addressing community development goals with limited resources

SALMONID ENHANCEMENT

➡ Goal

- Increase the number of salmon returning to their home rivers to a self-sustaining level

➡ Interventions constitute the main program theory elements

- Increase the number and improve the quality of potential breeding grounds in the river
- Ensure that river passages will allow salmon to move up and down the river
- Reduce harvesting through fisheries management and inhibiting poaching and overfishing
- Restock river with young fish

➡ Intervening factors

- Salmon spend a number of years at open sea
- Suffer unknown mortality at sea due to regulated and unregulated commercial harvesting, worsening environmental conditions, drift nets and other unintended human interventions

➡ Reach of the program interventions only extends to the fish departing the rivers

➡ Success requires that the fish return and reproduce successfully and is highly contingent on the intervening factors

ENVIRONMENTAL CONFLICT RESOLUTION

➡ Goal

- More effectively resolve a dispute over the use or quality of a resource

➡ Interventions

- Third party neutral facilitates parties to communicate with each other, narrow the issues to those that are central to the dispute and develop an agreement that is acceptable to all parties and which can endure
- Effectiveness depends directly on post agreement environmental and economic outcomes

➡ Reach of the program interventions only extend to the conclusion of an agreement

- Parties are responsible for implementation
- What parties do might not itself directly affect substantive environmental outcomes
- ADR practitioners also claim contributions to improved democracy as a substantive gain

➡ Evaluation challenges

- Attribution – what proportion of any substantive change can be attributed to ECR
- Incrementality – what difference does ECR make compared to a reasonable counterfactual
- Effectiveness – serious measurement issues

➡ Successful interventions are only marginally related to the goal

INTERNATIONAL PEACEMAKING/KEEPING AND CONFLICT RESOLUTION

➡ Goal

- return to sufficient stability and security for improved governance and development

➡ Interventions

- armed peace keepers, security sector (police, judiciary) reform, conflict resolution approaches, curtailment of small arms and light weapons trade

➡ Evaluation challenges

- Attribution – what proportion of any substantive change can be attributed to the intervention
- Incrementality – what difference did the intervention make compared to a reasonable counterfactual (e.g. armed peace keeping versus peacemaking, conflict resolution)
- Context – serious information gathering challenges, many parallel processes underway simultaneously and involving many of the same parties

➡ There is no agreed program theory, or even a candidate for one

URBAN ENVIRONMENTAL SERVICES

➡ Funder Goals

- Poverty reduction

➡ Interventions

- Latrines, pumped water from clean sources, passage (road and path) and safety improvements

➡ Outcomes that can be achieved with the interventions

- Reduce the incidence and seriousness of diseases attributable to sanitation and water
- Positive health effects can themselves be marginal
- Reduced incidence of health problems can reduce interruptions to income stream and the need to dispose of any assets (e.g. bicycle) for survival

➡ Examples of intervening factors

- Water piped to community distribution points and carried by hand to houses, stored in containers inside or outside the house
- Population open to contamination from external sources such as purchased food

➡ Longest reach of program interventions only extends to improved health

- Improved health has very marginal effect on household accumulation and poverty reduction

➡ Successful interventions are only marginally related to the goal

- Likely one of many necessary conditions

COMMUNITY DEVELOPMENT

➡ Community Development Goals

- Address serious economic problems of unemployment, low earnings, stagnant accumulation

➡ Interventions

- SME support, literacy and occupational training, modest investment funds for inward investment

➡ Outcomes that can be achieved with the interventions

- Marginal economic, capacity, and self esteem gains for program participants
- Modest investments in likely marginal enterprises

➡ Context

- Community development investments are a tiny fraction of the total economic development investment

➡ Effects of the community development interventions are too small to be noticeable

SUMMARY OF THE CHALLENGES

➡ Salmon Enhancement

- The interventions can greatly increase the fish leaving the rivers
- Success is highly contingent on parallel successful interventions at sea

➡ Environmental Conflict Resolution

- The interventions seeks to gain an agreement to settle an environmental dispute
- The interventions are not directly associated with anything to do with the environment

➡ Peacekeeping/making and Conflict Resolution

- There is no accepted program theory to work with

➡ Urban Environmental Improvement

- The interventions are marginally related to the goal of poverty reduction

➡ Community Development

- Effect size far too small

SOME OPTIONS TO EXTEND REACH OF INTERVENTIONS

Credible Options

1. Model outcomes outside the reach of the interventions based on sound program theory (e.g. simulation models)
2. Observe key outcomes at the interface of the program theories of the intervention and one that links (reaches to) the intended effects

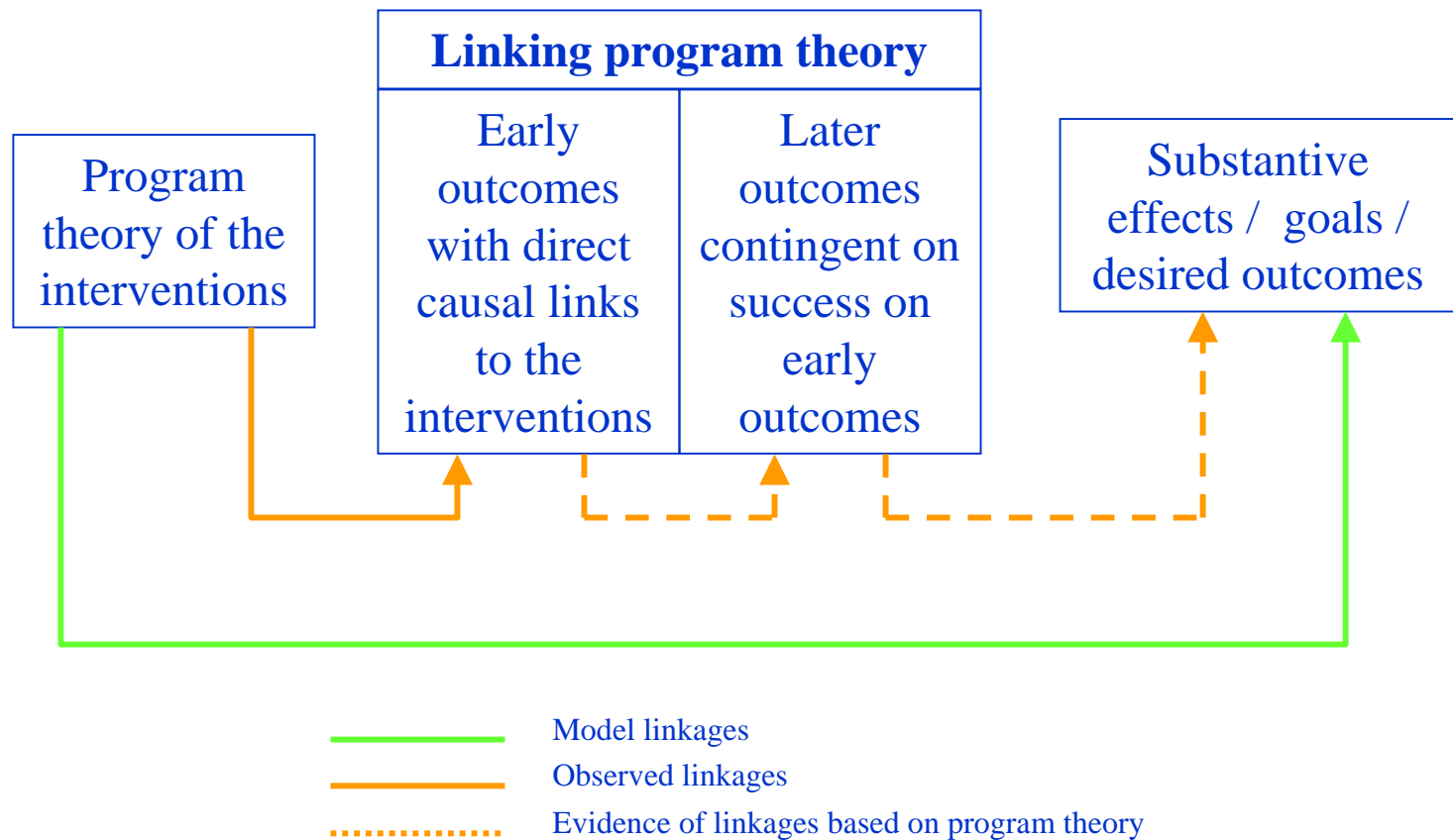
Not Infrequent, Not Very Credible Options

1. Model outcomes outside the reach of interventions based on a program theory with limited evidence base
2. Make claims at a high level of generalization to extend the reach but also reducing the plausibility or strength of impacts

OR

5. Claim to be an economist thereby lowering expectations of the reliability of forecasts

COMPARE MODEL TO OBSERVE



MODEL OR OBSERVE?

➡ Model Desired Effects

- Focuses on main desired effect such as return of salmon to the rivers
- Requires reliable evidence based predictors of the desired effects
- Can model effects of at least one leading counterfactual – doing nothing
- Difficult to model a full program theory
- Permits testing of a range of assumptions about intervening factors useful for risk management but not necessarily enhancing the reliability of the effect forecasts

➡ Observe Links to Desired Effects

- Focuses on intermediate observable outcomes in a program theory linking the interventions to the desired effects
- Requires a linking outcome based program theory
- Difficult to develop a similar structure for a likely counterfactual
- Difficult to make statements about the relative importance of the interventions to the desired effect
- Challenge gaining credibility amongst stakeholders because the main desired effect is not observed

EXAMPLE OF A MODELLING APPROACH

➔ Salmonid Enhancement

- Evaluation of the Canada/Newfoundland Salmonid Enhancement Agreement (circa 1996)
- Obtained evidence that the immediate effects of the interventions were being achieved (e.g. planted salmon surviving and leave river when intended, salmon clustering or using cleaned up passages and steadies) or not achieved (in-river and estuary poaching not reduced, no improved adherence to recreational fishing regulations, community watershed management no better than governmental management)
- Developed a simulation, inputs were the biological outcomes to the return of the salmon, output was key dependent variable salmon returns, with considerable front end user control over assumptions about human and environmental intervening variables
- Able to make conditional statements about the likely contribution of the main program interventions to desired effect and the parameters of the main effect (number of salmon returning) based on accepted good theory and evidence
- Also able to make statements credible to stakeholders about the counterfactual of no enhancement program

BUT HOW TO DO WE KNOW IF THE SCIENCE IS GOOD ENOUGH?

- ➡ **The salmon counts increased for the first two years the stocked salmon were at sea, then decreased.**
 - Counts of returning fish never reached sustaining levels
- ➡ **There are plenty of candidates for the role of nasty intervening variables**
 - Commercial harvesting in European and Russian waters or waters outside the 200 mile zones by south European and Asian fleets
 - Continuing degradation of the marine environment
 - Human created mortality factors such as lost-at-sea nylon gill nets drifting below the surface
 - Near-shore poaching
 - Cuts in funding for the fish counting and river monitoring programs
- ➡ **There is no good science about the intervening variables**

WHAT WAS THE EVALUATION ABLE TO SAY?

➡ Actual reporting period in year 4 of the initiative

- Program had been undertaken with fidelity to know good science
- Some program elements had not been as successful (enforcement and community watershed management)
- It was likely that the salmon would return but whether sustainable levels would be reached was unknown, and,
- Without the program the salmon would have been unlikely to return
- Overall, the enhancement program was successful

➡ In hindsight after the salmon did not return, then returned, then were gone again

- Was the program dosage off? Were continued funding and interventions necessary?
- Was the science off?
- Was enhancement the appropriate intervention? Would diplomatic initiatives have had better results?

➡ Are there evaluation approaches that would have been more useful to stakeholders

- Seems to require that the knowledge base for the evaluation is better than the existing science

ASSESSMENT OF MODELING TO GET TO LIKELY TO

➔ Many advantages

- If the science is good modeling should get to the results with sufficient quality
- Can model incrementality of the intervention through counterfactuals
- Efficient tool that can be used collaboratively enhancing utilization

➔ Challenges

- Hard to model intermediate outcomes upon which the desired outcomes are contingent – e.g. success of the salmon at sea, questionnaires administered to the salmon to gain insights into their intentions to return to the rivers
- Part of the need for the model is the long time period for the desired result to occur making modeling a high risk evaluation approach
- Without intermediate contingent outcomes evaluation cannot provide guidance to the program on the need for additional interventions to enhance likelihood of success

EXAMPLE OF A LIKELY TO APPROACH USING OBSERVATIONS

➡ Environmental Conflict Resolution

- Collaborative process facilitated by a third party neutral

➡ Environmental Conflict

- E.g. a fish passage dispute – dams operated by a local electric utility who generate power for sale and also sell the water to downstream farmers for irrigation but which also obstructs the passage of fish beyond the dams
- Native Americans have legal and ancestral rights and responsibilities for the health of the fish, and to harvest the fish
- Litigation or an administrative decision by government would be a win-lose, ECR seeks win-win

➡ Main evaluation issues

- Very difficult to gain access to a sufficiently similar counterfactual
- Substantive environmental outcomes such as fish health, power sales, success of alternative irrigation methods will take time to occur and can be difficult to observe

PROGRAM THEORY OF ECR AT THE US EPA CONFLICT PREVENTION AND RESOLUTION CENTER

Process Outcomes for Agreement Seeking and Non-Agreement Seeking Processes

Right parties are
effectively engaged

Appropriate
process scope and
design is used

Parties have
capacity to engage
in the process

Parties understand
issues/ narrow
areas of
disagreement to
key issues

Appropriate neutral(s)
guides the process

Best information (legal,
technical, etc) used by
parties

Parties communicate
and collaborate
effectively

Parties' capacity to
manage and resolve
conflicts is improved

All parties are satisfied
with the process

Agreement Outcomes for Agreement Seeking Processes

Issues that parties
cannot agree on
addressed with
other approaches

Parties reach
complete and
durable
agreements

All parties are
satisfied with what
they have achieved

Downstream Outcomes

Agreements are
implemented

Agreements endure
changes in
conditions and
unanticipated
events

Parties will use
collaborative
processes more
frequently and
expend fewer
resources on
disputes

Reduced frequency
and intensity of
disputes



CPRC Accountable

Shared
Accountability

THE GAP

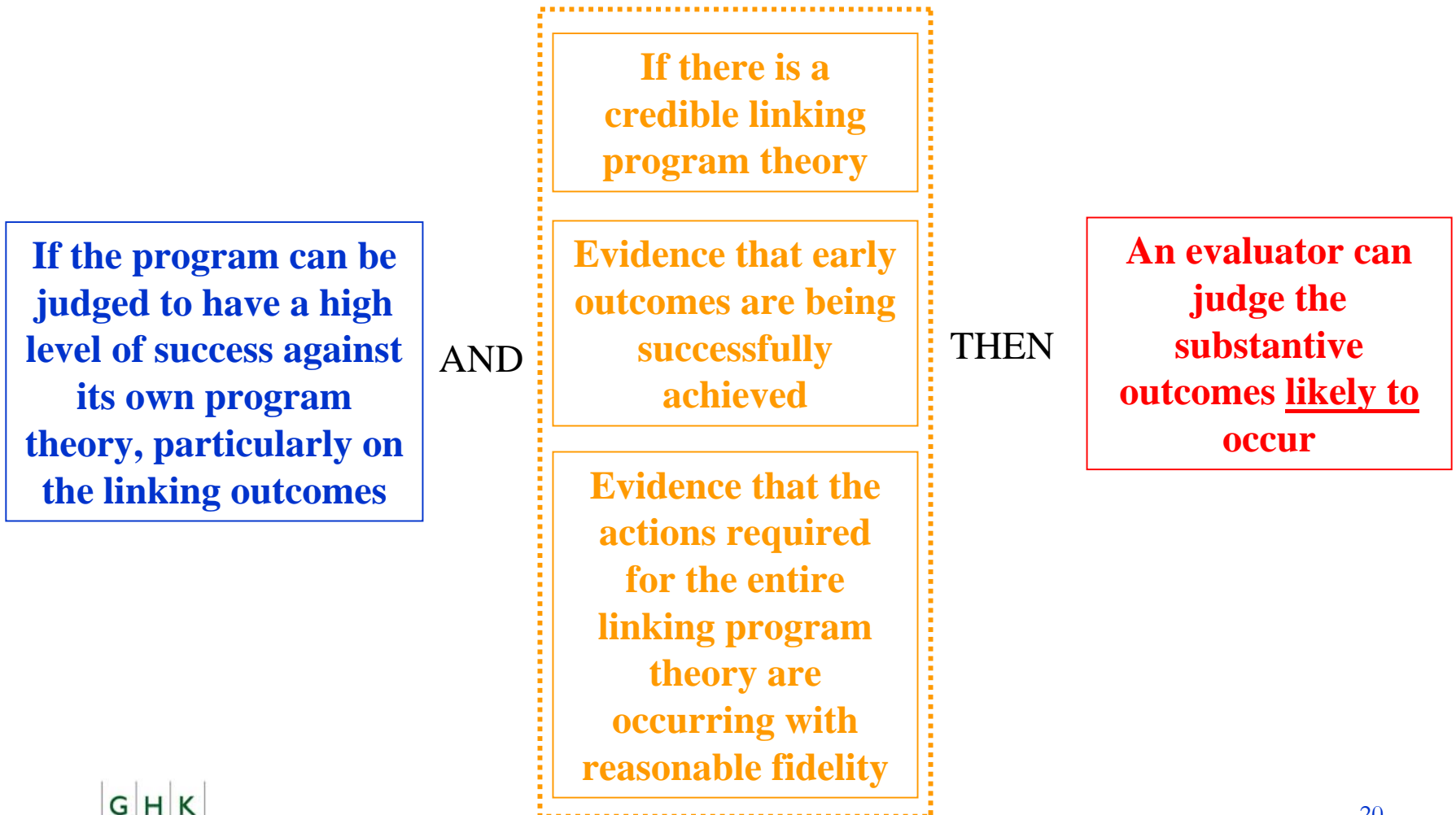
➡ The ECR program theory does not reach to environmental impacts

- Note the accountability line (blue) in the preceding ECR chart
- Outcomes to the right of this line are contingent upon the performance of others
- Environmental impacts lie even further to the right of the accountability line, are even more contingent

➡ Even when the program is well designed and implemented substantive effects can be well outside their reach

- Presumably some day when we have fully “joined up” government this will be much more exceptional than it is now

LINKING PROGRAM THEORY REQUIRED FOR LIKELY TO FROM OBSERVATION



REQUIRED CONDITIONS FOR AN ACCEPTABLE BRIDGING PROGRAM THEORY

➡ Requirements for the program under review

- The program theory of the program under review must be outcome based
- The performance of the program under review must be highly successful
- The validity and reliability of the measures used to assess performance on the program theory of the program under review must be established and sufficient

➡ Requirements for the linking program theory

- The linking program theory must be outcome based
- There is a credible direct link from the linking program theory to the desired substantive outcomes
- There is valid and reliable evidence that the early outcomes of the linking program theory are being achieved
- There are plausible indications that the other outcomes of the linking program theory are occurring, or no evidence that they are not occurring

TWO APPROACHES TO LIKELY TO HAVE BEEN SUGGESTED

➡ **Articulate the entire program theory in a simulation model**

- Forecast results for substantive outcome based on performance of the program
- Advantage – can reach to substantive outcomes when it is not feasible to observe outcomes reaching from the program to the substantive outcomes
- Constraint – relies very heavily on the quality of the knowledge underlying the program theories
- Downside – evaluation cannot provide interim guidance for program improvement

➡ **Articulate a linking program theory and observe performance beyond the reach of the program**

- Observe levels of performance on contingent outcomes linking the program's own program theory to the substantive outcomes
- Advantage – evaluation can provide guidance for program improvement
- Constraint – accountabilities of stakeholders in the program often do not reach to the linking outcomes
- Downside – credibility is impaired because there is no direct observation of desired substantive outcome

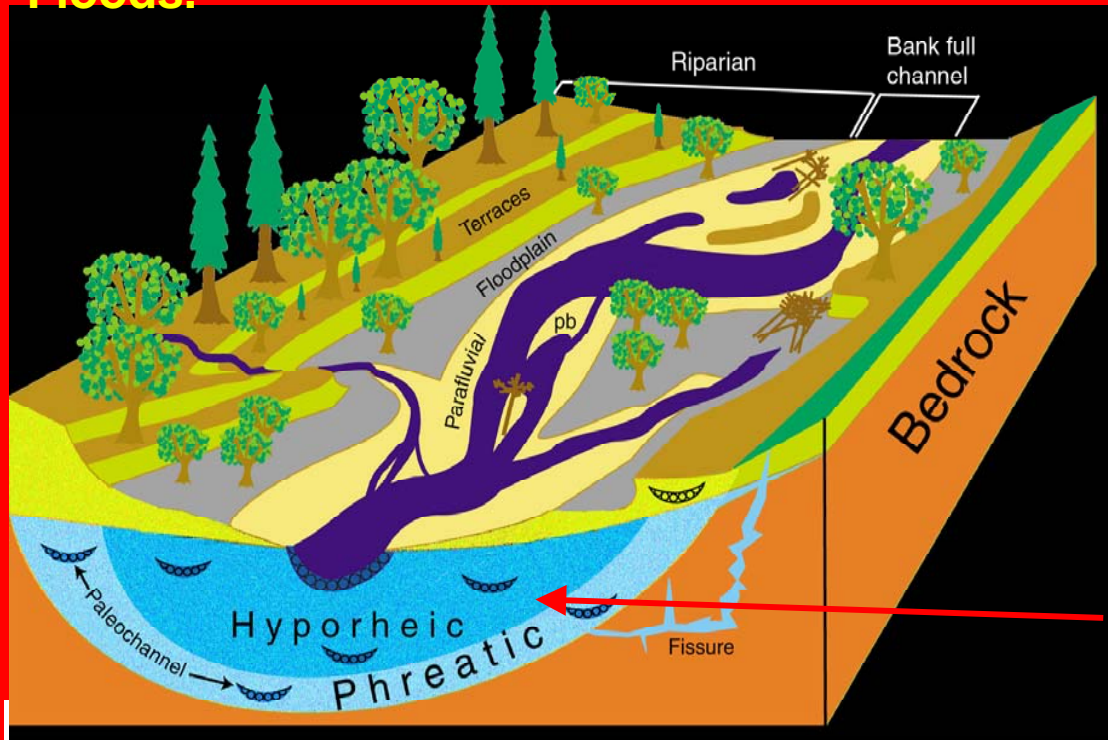
IS LIKELY TO GOOD ENOUGH?

- ➔ No – but what are the alternatives?
- ➔ Governance is moving slowly to results based accountability
 - But program silos still limit the reach of programs
 - Accountability and results focus is still embryonic
- ➔ While the utility and validity of the use of “program theory” in evaluation is now accepted by evaluators
 - The knowledge base on which program theories are built is still in silos
- ➔ Still, likely to seems to be a workable option for evaluation while science and governance catch up

SHELF LIFE OF LIKELY TO WILL EXPIRE WITH BETTER SCIENCE

- ➡ **Likely to will do until we have more dynamic complex knowledge base for program theories**
 - Evaluation can then assess these programs with our usual questions with our usual methods
- ➡ **Example: Modeling flood flows in Rocky Mountain rivers (Mark Lorang et al, University of Montana)**
 - Dynamic, historical, evidence based complex system models of river basins and effects of floods
 - This type of science should cause a rethinking of established wisdom of key environmental and resource management approaches
 - The models have a proven validity and reliability, are comprehensive, dynamic and embrace the complexity of ecosystems
 - Final and intermediate substantive outcomes resulting from different water management approaches can be modeled accurately enough for immediate water management decisions AND the environmental consequences including effects on species living under the river bottoms, endangered species in the river systems, etc.
- ➡ **Where we do not have this science base, likely to is a reasonable option for evaluation**

- Alluvial floodplains are BioDiversity and Production “hotspots”.
- The Shifting Habitat Mosaic (above and below the surface) of floodplains is maintained by the dynamics of cut and fill alluviation and avulsion.
- Biogeochemical cycling, system BioProduction and BioDiversity are directly linked to the dynamics of the Shifting Habitat Mosaic.
- Maintaining the natural Shifting Habitat Mosaic is dependent upon Floods.



CURRENT END

➔ Remainder are working slides

WHERE SCIENCE IS CURRENTLY GOOD ENOUGH – MINERAL DEVELOPMENT

CRYPTIC DESCRIPTION OF THE LOGIC OF A PROGRAM THEORY

If we have good geodesic maps

We will know how much to drill and where

We know the likely payback to exploration

And the environmental effects of the mines

And what this will mean to the companies, households, communities, regions, industries and nations

CRYPTIC DESCRIPTION OF THE MAIN INPUTS TO A MODEL

Geodesic mapping enables exploration firms to reliably site exploration drilling

For most minerals we know the number of exploratory holes that will be drilled given the results of the geodesic surveys

We know the coefficients linking the number of mines developed to the number of exploratory holes drilled

We know the economic effects of developing a mine

ILLUSTRATION USING ECR

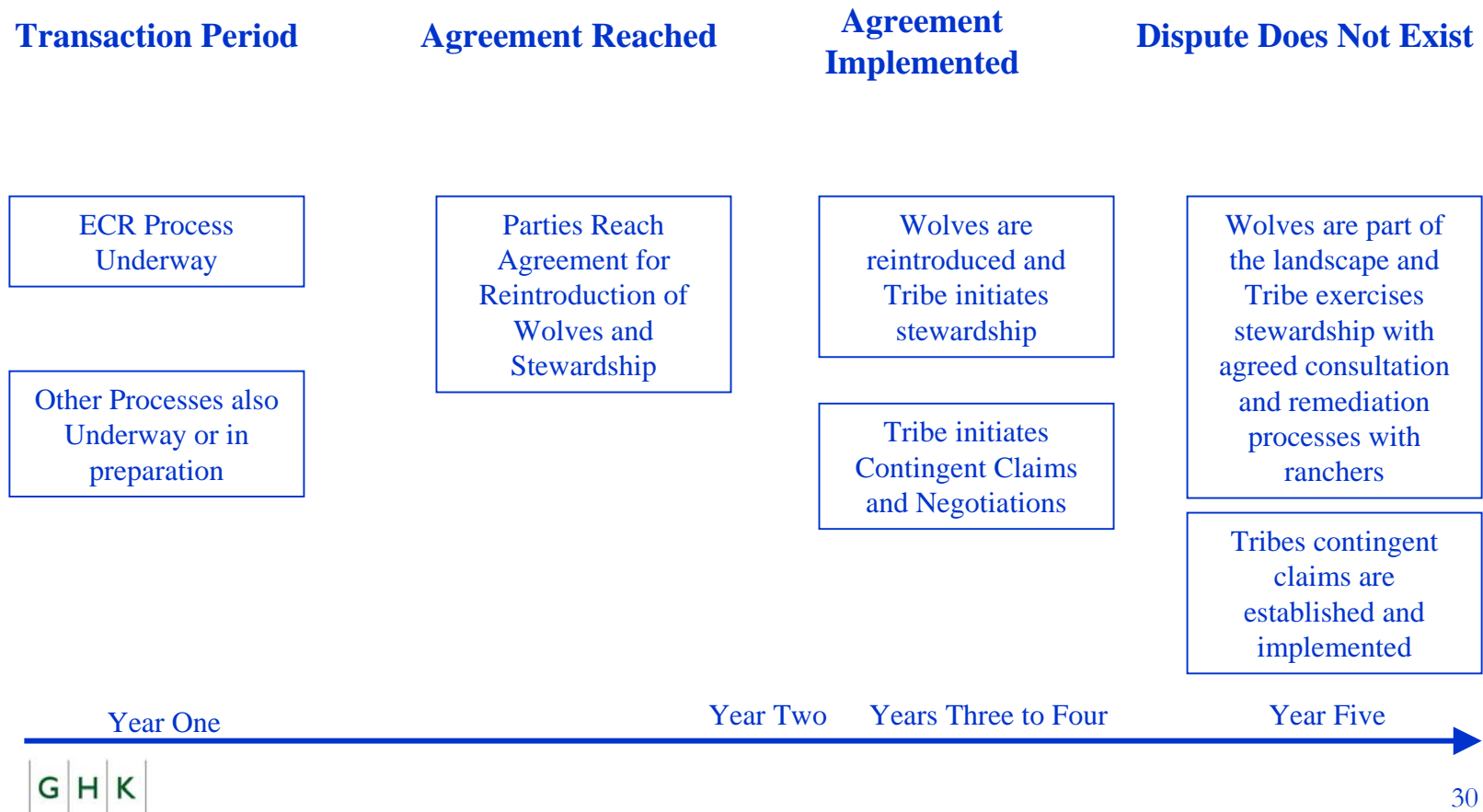
1. **The ECR program theory is outcome based**
 1. Uses the authors *outcome charting* approach
2. **The ECR program theory has a high level of face validity**
 - i. Frequent positive peer reviews by leading academics and practitioners in the field
 - ii. Willingness of additional federal and state programs to adopt the program theory with minor modifications
3. **The ECR program theory and its measures have been shown to be both valid and reliable**
 - i. Initial testing by developer confirmed validity and reliability
 - ii. Recent additional testing by independent statisticians confirmed validity and reliability
 - iii. Response rates to questionnaires for most agencies range from 74% to 94% (one agency has low response rates so their results would not pass this test) and there are no identifiable response or item response problems
 - iv. Initial evaluation findings are used by agencies and practitioners to make program and practice decisions and revise approaches to professional development

SUBSTANTIVE OUTCOMES AND LINKING PROGRAM THEORY

Nature of the dispute

1. A Native American tribe has established rights to stewardship of wildlife in an area, but their claim to the land on which the wildlife reside is not recognized by US government or US courts
2. Using collaborative processes an agreement is reached to reintroduce wolves to the area. Wolves have been absent over a hundred years but prior to that had a long established residency in the area
3. Parties to the agreement include state government, federal agencies, local farmers and ranchers, and the Tribe
4. With the restoration of the wolf population and stewardship of the Tribe over the wolf, the Tribe gained contingent rights including gaming, land tenure and other quasi Treaty rights
5. Other processes including preparation for litigation initiated by the Tribe were underway during and parallel to the collaborative processes

GUESSING AT THE LINKING PROGRAM THEORY



OUTCOME CHARTS ARE DESIGNED TO ADDRESS THESE THREE ACCOUNTABILITY QUESTIONS

1. What is my agency / business unit supposed to achieve is answered through the process and result of developing the outcome charts
2. How will we recognize it when it happens is answered by ensuring that all of the outcomes are observable
3. How are we doing now is addressed with the outcome monitoring system based on the outcome charts